

Experimental detection of entanglement with optimal-witness families

Jibo Dai,^{1,2} Yink Loong Len,^{1,2,3} Yong Siah Teo,^{2,4} Berthold-Georg Englert,^{2,3} and Leonid A. Krivitsky¹

¹Data Storage Institute, Agency for Science, Technology and Research, 5 Engineering Drive 1, Singapore 117608, Singapore

²Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore

³Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542, Singapore

⁴Department of Optics, Palacký University, 17. listopadu 12, 77146 Olomouc, Czech Republic

(Posted on the arXiv on 28 March 2014)

We report an experiment in which one determines, with least tomographic effort, whether an unknown two-photon polarization state is entangled or separable. The method measures whole families of optimal entanglement witnesses. We introduce adaptive measurement schemes that greatly speed up the entanglement detection. The experiments are performed on states of different ranks, and we find good agreement with results from computer simulations.

PACS numbers: 03.65.Ud, 03.65.Wj, 03.67.Mn

Introduction Entangled states play an important role in the manipulation of quantum information, be it for present-day quantum key distribution or future quantum computation. One may need to verify if a certain quantum state—perhaps emitted by a source of quantum-information carriers or obtained as the output of a quantum computation—is entangled or not. For this purpose, the expectation value of an *entanglement witness* is telling: The state is surely entangled if a negative value is obtained. A positive value, however, is inconclusive—the unknown state could be entangled or separable, the witness cannot tell, but other witnesses might be able to. The concept of a witness was first used by M., P., and R. Horodecki [1], and the term “witness” was introduced by Terhal [2]; for reviews that cover all important aspects of entanglement witnesses, see Refs. [3, 4].

How many witnesses, then, does one need to measure until a conclusion is reached? The answer to this question is given in Ref. [5]: If one solely relies on the expectation values of the witnesses one by one, one may never get a conclusive answer; if, however, the expectation values of suitably chosen witnesses are jointly used for an estimation of the quantum state, $D^2 - 1$ witnesses suffice for a D -dimensional quantum system.

The number can be further reduced by exploiting all the information gathered when determining the expectation values of the witnesses measured in succession. In the case of a two-qubit state ($D = 4$), for instance, one never has to measure more than six witnesses, rather than $15 = 4^2 - 1$. This can be demonstrated by an experiment such as the one proposed in Ref. [5]; we are here reporting its laboratory realization.

Witnesses and witness families A hermitian observable W is an entanglement witness if $\text{tr}\{\rho_{\text{sep}} W\} \geq 0$ for all separable states ρ_{sep} and $\text{tr}\{\rho_{\text{ent}} W\} < 0$ for at least one entangled state ρ_{ent} . For each entangled state, there are some witnesses that detect it (“ < 0 ”), but many other witnesses will give an inconclusive result (“ ≥ 0 ”).

We shall concern ourselves with two-qubit systems—in the experiment, they are polarization qubits of a down-converted photon pair—and focus on optimal decomposable witnesses [6] of the form $W = (|w\rangle\langle w|)^{T_2}$, where $|w\rangle$ is the ket of

an entangled pure two-qubit state, and T_2 denotes the partial transposition on the second qubit. This witness is optimal in the sense that no other witness can detect some entangled states in addition to the states already detected by W . The generic example is $|w\rangle = |00\rangle \cos(\frac{1}{2}\alpha) + |11\rangle \sin(\frac{1}{2}\alpha)$ with $\sin \alpha \neq 0$, on which all other $|w\rangle$ s can be mapped by local unitary transformations.

For any α , the eigenkets of the resulting witness,

$$W^{(\alpha)} = |00\rangle \frac{1 + \cos \alpha}{2} \langle 00| + |11\rangle \frac{1 - \cos \alpha}{2} \langle 11| + |\Psi_+\rangle \frac{\sin \alpha}{2} \langle \Psi_+| - |\Psi_-\rangle \frac{\sin \alpha}{2} \langle \Psi_-|, \quad (1)$$

are the same: the two product kets $|00\rangle$ and $|11\rangle$ as well as the two Bell kets $|\Psi_{\pm}\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$. This whole family of entanglement witnesses can, therefore, be measured by the projective measurement of their common eigenstates—the *witness basis* of the family. For a separable state, $\text{tr}\{\rho_{\text{sep}} W^{(\alpha)}\} \geq 0$ for all α , and this requirement implies the witness-family criterion [5]

$$\mathcal{S} \equiv 4f_1 f_2 - (f_3 - f_4)^2 \geq 0. \quad (2)$$

Here, f_1 and f_2 are the probabilities for the two product states, and f_3 and f_4 are those for the Bell states; we estimate these probabilities from the observed frequencies. Consequently, once the frequency data are obtained from the measurement, a negative value of \mathcal{S} reveals that the unknown state ρ_{true} is entangled.

Two remarks are in order. (i) The expectation value of one witness $W^{(\alpha)}$ is a linear function of the f_j s, whereas \mathcal{S} is a quadratic function. This is reminiscent of, yet different from, the “nonlinear entanglement witness” of Ref. [7], which requires a joint measurement on two copies of the unknown state. Our *witness-family measurement* uses only one copy at a time. (ii) The witnesses $W^{(\alpha)}$ can also be measured by other schemes, such as that of Barbieri *et al.* [8] who extracted the expectation value of $W^{(\pi/2)}$ from local measurements that examine the two qubits individually. By contrast, we perform a joint measurement on both qubits, thereby realizing the most direct measurement of $W^{(\alpha)}$.

TABLE I: The six witness families that enable full tomography of the two-qubit state. The single-qubit unitary operators U_1 and U_2 transform the first family into the other five families. The Pauli operator X permutes $|0\rangle$ and $|1\rangle$; the Clifford operator C permutes the three Pauli operators cyclically.

Witness	U_1	U_2	Witness	U_1	U_2
1	$\mathbf{1}$	$\mathbf{1}$	4	C^\dagger	XC
2	$\mathbf{1}$	X	5	C	C^\dagger
3	C^\dagger	C	6	C	XC^\dagger

Our witness-family measurement provides estimates for *three* two-qubit observables (four f_j s with unit sum), whereas the expectation value of a single witness is only one number. As discussed in Ref. [5], this can be exploited for quantum-state reconstruction after measuring six witness families, related to each other by the six local unitary transformations of Table I. A measurement of all six witness families constitutes an informationally complete (IC) measurement for full tomography of an unknown two-qubit state. This offers the possibility of measuring an IC set of witness families such that, if all families give an inconclusive result ($\mathcal{S} \geq 0$), a full state estimation can be performed for identifying ρ_{true} . With ρ_{true} then at hand, its separability can be determined straightforwardly by, for example, checking the Peres-Horodecki criterion [1, 9].

Scheme A: Random sequence Since ρ_{true} is unknown, there is no preference for a particular one of the six families to start with. Hence, one can measure the six families in a pre-chosen order or in a random sequence. In this scheme, one starts with a randomly chosen family and checks the inequality (2). If the result is inconclusive, one then chooses the next family at random from the remaining five families, and so forth until a conclusive result is obtained. If all six families give inconclusive results, ρ_{true} is estimated from the data to establish if it is entangled or separable.

Scheme B: Adaptive measurements Alternatively, we can perform the witness-family measurements in an adaptive manner: We choose the next family to be measured in accordance with the data obtained from previous measurements. Each time a witness family is measured, a set of four frequencies is obtained and these *informationally incomplete* data are used to partially estimate ρ_{true} by, for example, jointly maximizing the likelihood and the entropy—the MLME strategy of Ref. [10, 11]. The MLME estimators ρ_{MLME} tend to be highly-mixed states and are thus hard to detect by entanglement witnesses; therefore, if the measurement of a witness family detects the entanglement of the MLME estimator by the criterion (2), measuring that family has a good chance of detecting the entanglement of ρ_{true} . The value of \mathcal{S} is used for comparing the unmeasured witness families with the MLME estimator, with f_j replaced by the j th MLME probability $\text{tr}\{\rho_{\text{MLME}} U^\dagger \Pi_j U\}$ where $U = U_1 \otimes U_2$ is one of the six unitary operators of Table I, and Π_j projects to the j th ket in the witness bases. The family that gives the smallest

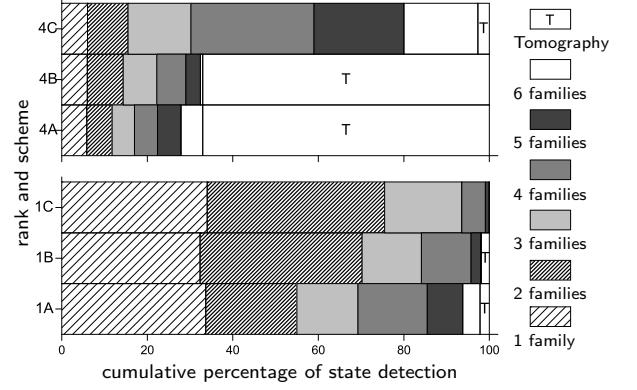


FIG. 1: Simulation results on the measurement of the set of six informationally complete entanglement witness families for 10^4 randomly chosen two-qubit entangled states: pure states (bottom) and full-rank mixed states (top). The cumulative histograms compare between measurements performed with scheme A, scheme B, and scheme C.

value of \mathcal{S} is measured next; this judicious choice of family reduces the average number of witness families that need to be measured before the entanglement is detected. Instead of fixing the above six families, based on the MLME estimator, one can also choose from all thinkable families in each step. However, it is not worth the trouble as such optimization hardly improves the entanglement detection; see the Appendix for details.

Scheme C: Maximum-likelihood set All state estimators, including ρ_{MLME} , that maximize the likelihood compose a convex set, the *maximum-likelihood (ML) set*. They all give the same estimated probabilities for the witness families already measured. When the number of qubit pairs measured per witness family is large (10^4 pairs suffice in practice), ρ_{true} is very likely contained in the ML set. Then, if there is no separable state in the ML set, we can conclude that ρ_{true} is entangled. For finite data, this conclusion is correct within a certain error margin, as is the case for all conclusions drawn from entanglement witness measurements. For this non-separability check, we compute the maximal values of the likelihood for both the entire state space (\mathcal{L}_{max}) and the entire space of separable states ($\mathcal{L}_{\text{max}}^{\text{sep}}$) [12]. If we find that $\mathcal{L}_{\text{max}} > \mathcal{L}_{\text{max}}^{\text{sep}}$, we infer that ρ_{true} is entangled. To further economize the adaptive scheme, this check is performed before looking for the unmeasured witness family with the smallest value of \mathcal{S} .

Simulations To investigate the efficiencies of the three schemes, we perform computer simulations of witness-family measurements with both pure and full-rank mixed two-qubit entangled states. Figure 1 shows the histograms that summarize the cumulative distribution in the percentage of entangled states detected against the number of witness families needed using schemes A, B and C. We observe that the average number of families is largest for scheme A and smallest for scheme C. When using either scheme A or scheme B,

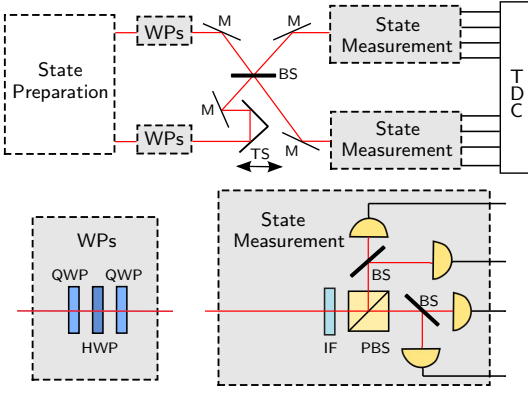


FIG. 2: (color online). Experimental set-up. The polarization-entangled two-photon states are prepared by the method described in Ref. [13]. Upon emerging from the source, the two photons are guided with mirrors (M) to interfere at a 50:50 beam splitter (BS), with the temporal overlap controlled by a translation stage (TS). After passing through interference filters (IF), the photons are sorted by polarizing beam splitters (PBS) and registered by one of the photo-detectors, four on each side. The detector outputs are addressed to a time-to-digital converter (TDC), and coincidences between counts of any two detectors are recorded. Two sets of wave plates (WPs), each composed of a half-wave plate (HWP) and two quarter-wave plates (QWP), implement the polarization changes that correspond to the unitary operators of Table I.

about 2% of the random pure states and about 67% of the full-rank mixed states are undetected by the six families without performing full tomography. The additional separability check in scheme C reduces the percentage of undetected pure states to virtually zero, and one needs no more than five witness families to detect entanglement for the rest of the pure states. The improvement is even more dramatic for the mixed states, with a reduction from about 67% to about 2.7%. We also observe that the mean number of witness families needed to detect entanglement for mixed states is larger than that for pure states. This is as expected, since mixed states generally have weaker entanglement and are, therefore, harder to detect.

Experiment We experimentally test the entanglement detection and tomographic scheme with three classes of states of different ranks. The first class of states are the pure states $\rho_{\text{true}}^{(1)} = |\vartheta\rangle\langle\vartheta|$, with $|\vartheta\rangle = |00\rangle \sin \vartheta + |11\rangle \cos \vartheta$ for $0 < \vartheta < \pi$, $\vartheta \neq \pi/2$. The second class of states are rank-two states of the form $\rho_{\text{true}}^{(2)} = |\Phi_+\rangle\mu\langle\Phi_+| + |\Phi_-\rangle(1-\mu)\langle\Phi_-|$ for $0 \leq \mu \leq 1$ and $\mu \neq 1/2$, where $|\Phi_{\pm}\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$. The third class of states are the Werner states $\rho_{\text{true}}^{(3)} = |\Psi_-\rangle\lambda\langle\Psi_-| + (1-\lambda)/4$ for $1/3 < \lambda \leq 1$. The experiment uses the polarization qubits of a down-converted photon pair with, for example, ket $|10\rangle$ standing for the photon in mode 1 horizontally polarized and the photon in mode 2 vertically polarized.

The experimental set-up is shown in Fig. 2. For the state preparation, we use two type-I β -barium borate (BBO) crystals of 0.8 mm thickness with orthogonal optical axes [14]. The BBOs are pumped by a continuous-wave diode laser at

405 nm (Ondax Inc.), and their optical axes are orientated such that the down-converted photons at 810 nm are emitted at angles of $\pm 3^\circ$ to the direction of the pump. The rank-one states are produced by setting the polarization of the pump with a half-wave plate to an angle of $\vartheta/2$. The rank-two states are produced by dynamically switching the polarization of the pump using a liquid crystal retarder (LCR) [13]. Different states are obtained by varying the duty cycle of the LCR. The rank-four states are produced by adding a controlled admixture of white noise to the singlet state by varying the coincidence time window of the detection electronics [15].

The signal and idler photons are directed to a Hong-Ou-Mandel (HOM [16]) interferometer with a 50:50 beam splitter (BS), whereby a translation stage with a step size of 500 nm optimizes the temporal overlap between the photons. Typical obtained visibilities of the HOM dip exceed 90%. On the way from the source to the BS, the photons pass through sets of wave plates (WPs) that change the polarization in accordance with one of the six local unitary transformations of Table I. See the Appendix for details of the HOM calibration and the setting of the WPs.

In each output port of the interferometer, the photons pass through an interference filter (IF) with a central wavelength at 810 nm and a full width at half maximum of 10 nm and are then sorted by a polarizing beam splitter (PBS). To discriminate between one-photon and two-photon events, another 50:50 BS is installed into each output port of the PBS, and eight single-photon avalanche photodiodes (APD, quantum efficiency $\sim 30\%$, qtools GmbH) detect the photons. A time-to-digital converter (TDC, qtools GmbH) records the arrival times of the photons, and coincidences between any two of the eight detectors are obtained from the analysis of the time stamp record of the TDC. For each rank, we studied 21 different states.

Results We set the source to a particular state and performed witness-family measurements for one minute per family, and so measured about 10^4 photon pairs for each state and family. The data were analyzed for all three schemes.

We also performed simulations for the three different classes of two-qubit states. For each class, data from 10^4 photon pairs (per witness family) for 10^3 states were simulated with Monte Carlo techniques, for the schemes A, B, and C. As expected, the simulation showed that for measurements done with the random order of scheme A, the number of families needed for entanglement detection is distributed almost evenly from one to six, with a mean of about 3.5, and fewer families need to be measured in the adaptive schemes.

This is confirmed by the experiment. For the three classes of states, only one of the witness families gives a conclusive result. Hence, if the sequence of the six families were chosen randomly, the number of families needed for entanglement detection is equally likely to be one to six. The data for schemes B and C were analyzed as explained in Table II and the results are shown in Fig. 3, where one observes a significant improvement over the non-adaptive scheme A, which needs about 3.5 families on average. One rarely needs more than four wit-

TABLE II: Examples demonstrating how Fig. 3 is derived. The first family is chosen at random among the six families. If the measurement of this family gives a negative value of \mathcal{S} , the state is detected to be entangled and no further measurement is needed. If the result is inconclusive, then one uses the adaptive scheme to choose the next family, until a conclusive result is obtained. The result is n , the number of witness families that have to be measured in order to detect the entanglement.

State	Families and their \mathcal{S} values			n
$\rho_{\text{true}}^{(2)}$	family 3	family 4	family 2	3
$\mu = 0.15$	0.46 ± 0.03	0.06 ± 0.02	-0.23 ± 0.04	
$\rho_{\text{true}}^{(3)}$	family 1	—	—	1
$\lambda = 1$	-0.83 ± 0.03			
$\rho_{\text{true}}^{(1)}$	family 4	family 2	—	2
$\vartheta = \pi/4$	0.21 ± 0.03	-0.87 ± 0.03		

ness families to detect the entanglement in these states; full tomography is never necessary since all rank-one, rank-two, and rank-four states are detected by one of the witness families.

The striking similarity between the results of the simulation and the experiment indicates that there are no significant systematic errors in the experimental data. For the rank-one states, the fidelity $F = \sum_j (p_j p'_j)^{1/2}$ between two probability distributions $\{p_j\}$ and $\{p'_j\}$ compares the histograms from simulation and experiment, and the large value of F is reassuring. The error bars for the rank-one states are obtained by bootstrapping the actual data one hundred times; these error bars show the variation in the histograms that repeated measurements of this kind would display. The fidelity values and the error bars for the other histograms are of similar sizes and not displayed.

Tomography For the three classes of states, all entangled states are successfully detected without the need to perform full tomography. There are, however, other entangled states that would escape detection, and all separable states can only give inconclusive results. Regarding entangled states, we recall that about 2% of the random pure states and about 67% of the random mixed states are not detected by the six witness families without the separability check of scheme C. Hence, to confirm the efficiency and accuracy of the tomographic scheme, we collect IC data even if the state is known to be entangled before all six witness-family measurements are done. Using the technique of ML estimation [17, 18], we then infer the state from the data and calculate the fidelity $\text{tr}\{|\sqrt{\rho_{\text{true}}}\sqrt{\rho_{\text{est}}}|^2\}$ between the true and the estimated state. The average fidelities are $98.3 \pm 0.7\%$, $97.4 \pm 1.4\%$, and $98.7 \pm 1.1\%$ for the respective true states of rank one, two, and four. Indeed, if all the six family measurements fail to detect the entanglement in the state, one can use the tomographic information to reliably reconstruct the unknown input state and then determine its separability numerically.

Conclusions We performed an experiment to verify the

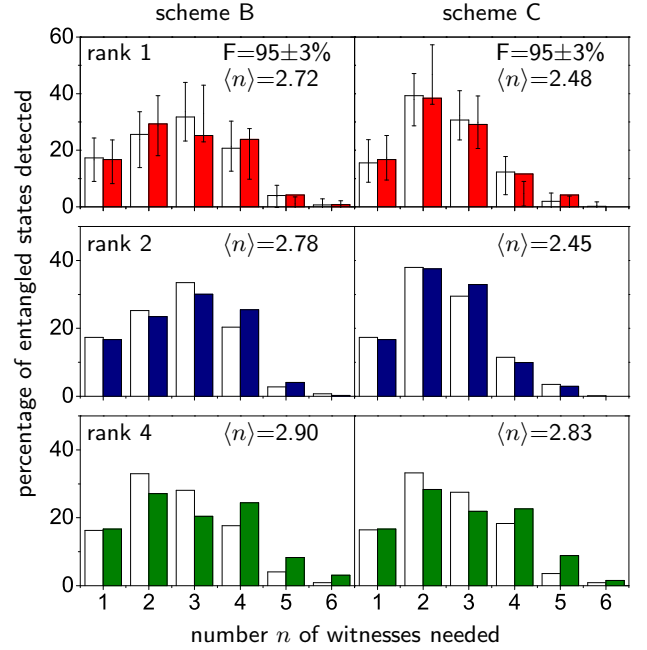


FIG. 3: (color online). A comparison of schemes B (left column) and C (right column) for rank-one states (top row), rank-two states (middle row), and rank-four states (bottom row). The histograms report the percentage of entangled states detected against the number n of witness families needed *without* performing state estimation; $\langle n \rangle$ is the average value. Both the simulation data (left empty bars) and the experimental data (right full bars) show that, for the three kinds of quantum states considered, scheme C provides further improvement over scheme B: It requires fewer families on average and the distributions are narrower. The similarity of the two histograms for the rank-one states is confirmed by their large fidelity F ; similar values are obtained for the other histograms.

witness-family-based entanglement detection scheme introduced in Ref. [5]. In going beyond that proposal, we also introduced adaptive schemes that use the information acquired in previous measurements to reduce the average number of witness families that need to be measured. An efficient way of detecting entanglement was thus demonstrated. Further, we showed that the witness-family measurements enable reliable quantum state tomography. The witness-family approach is also applicable to qubits of other physical kinds than photon polarization and so offers an efficient and versatile tool for entanglement detection.

Acknowledgments We are grateful for the insightful discussions with J. Řeháček and H. Zhu. This work is supported by the National Research Foundation and the Ministry of Education, Singapore, and is co-financed by the European Social Fund and the state budget of the Czech Republic, project No. CZ.1.07/2.3.00/30.0004 (POST-UP).

APPENDIX

In this appendix, we discuss more general adaptive schemes and provide some technical details of the experiment.

GENERAL ADAPTIVE SCHEMES

The adaptive schemes B and C that speed up the entanglement detection select the next witness family from the six pre-chosen families specified by the unitary operators in Table I. If, instead, one selects also from other families than the six pre-chosen ones, such a more general adaptive scheme might be more efficient, in the sense that fewer families need to be measured on average before one can conclude that the unknown state is entangled. Thereby, the selection is still done by opting for the family which is expected to give the smallest value of \mathcal{S} upon measurement. It turns out that the more general adaptive schemes are not worth the trouble. We justify this remark by a study of the generalizations of schemes B and C.

Scheme B' As in scheme B, we calculate the MLME estimator ρ_{MLME} and exploit its properties when choosing the next family to be measured. If ρ_{MLME} is entangled, $\rho_{\text{MLME}}^{\text{T}_2}$ has one negative eigenvalue, and the eigenket $|\phi\rangle$ to this eigenvalue is entangled. The next witness family is then the one obtained from $W = (|\phi\rangle\langle\phi|)^{\text{T}_2}$ because this family is best for detecting the entanglement of ρ_{MLME} . Since ρ_{MLME} is the current best guess for the unknown state ρ_{true} , this family has also a good chance of detecting entanglement in ρ_{true} . If the MLME estimator is separable, however, we proceed as in scheme B.

Scheme C' On top of scheme B', the separability check of scheme C is implemented.

Simulations We investigate the general adaptive schemes by performing computer simulations for pure and full-rank states and constructing histograms analogous to those in Fig. 1. As can be seen from the results in Fig. 4, scheme B' improves over scheme B. For both rank-one and rank-four states, the cumulative percentage of states detected is higher. The percentage of undetected rank-four states after six families drops from about 67% to about 25%. While this improvement is substantial, it pales in comparison with the dramatic reduction to about 2.7% when using scheme C. Therefore, when aiming at the most efficient way of detecting entanglement, we have to employ either scheme C or scheme C'.

Now, as we learn from Fig. 4, scheme C' is only slightly better than scheme C — if at all. For, the somewhat smaller proportion of rank-four states detected by the first, randomly chosen, family and the somewhat larger proportion of states requiring tomography are surely resulting from statistical fluctuations in the simulation. The other differences between the histograms for schemes C and C' are of similar size. Accordingly, there is no evidence that scheme C' is worth the trouble of its implementation, which requires that the settings of the WPs are calculated and updated in every step for every state. On the other hand, the data obtained in scheme C are opti-



FIG. 4: Simulation results for 10^4 randomly chosen two-qubit entangled states: pure states (bottom) and full-rank mixed states (top). The cumulative histograms compare between adaptive measurements performed with the six pre-chosen families of Table I (schemes B and C) and with six arbitrary families (schemes B' and C').

mal for the tomographic reconstruction of ρ_{true} . In summary, then, scheme C serves all purposes very well.

DETAILS OF THE EXPERIMENT

Wave plate settings Each of the local unitary operators of Table I is implemented by a HWP between two QWPs,

$$U = U_{\text{QWP}}(\alpha)U_{\text{HWP}}(\beta)U_{\text{QWP}}(\gamma), \quad (3)$$

where the QWP with angle γ is the first in the sequence; see Fig. 2. The matrix representation for a HWP, apart from an irrelevant global phase factor, is (see, e.g., [19])

$$\begin{aligned} U_{\text{HWP}}(\theta) &= \begin{pmatrix} |V\rangle & |H\rangle \end{pmatrix} \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \begin{pmatrix} \langle V| \\ \langle H| \end{pmatrix} \\ &\equiv \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}, \end{aligned} \quad (4)$$

where θ is the angle between its major axis and the vertical direction, and we have

$$U_{\text{QWP}}(\theta) \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 - i\cos(2\theta) & -i\sin(2\theta) \\ -i\sin(2\theta) & 1 + i\cos(2\theta) \end{pmatrix} \quad (5)$$

for a QWP. Further, the matrices for the Pauli operator X that permutes $|V\rangle$ and $|H\rangle$ and the Clifford operator C that permutes the three Pauli operators cyclically are

$$X \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad C \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}, \quad (6)$$

respectively.

TABLE III: Wave plate settings for the unitary operators of Table I. The angles α , β , and γ are such that the corresponding U is obtained from Eq. (3).

U	α	β	γ
$\mathbf{1}$	0	0	0
X	0	$\pi/4$	0
C	0	$\pi/4$	$-\pi/4$
C^\dagger	$-\pi/4$	0	0
XC	0	0	$-\pi/4$
XC^\dagger	$\pi/4$	0	0

The angles α , β , and γ , for which the various U s are realized, are reported in Table III. As an example, we consider

$$U = XC \hat{=} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \quad (7)$$

and verify that $\alpha = 0$, $\beta = 0$, and $\gamma = -\pi/4$ are correct choices. Indeed, they are:

$$\begin{aligned} & U_{\text{QWP}}(0) U_{\text{HWP}}(0) U_{\text{QWP}}(-\pi/4) \\ & \hat{=} \frac{1-i}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & i \end{pmatrix} \\ & = e^{-i\pi/4} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}, \end{aligned} \quad (8)$$

since the global phase factor is irrelevant.

HOM interferometer To implement the most direct measurement of the witness bases, we make use of a HOM interferometer. For its optimization, we first remove the BSs in the output ports of the PBSs; see Fig. 2. Next, we address the outputs of the two detectors in the transmission ports of the two PBSs to a coincidence unit, where we record the coincidence rate for the input two-photon state $|\text{HH}\rangle$. A translation stage with a step size of 500 nm is used to control the

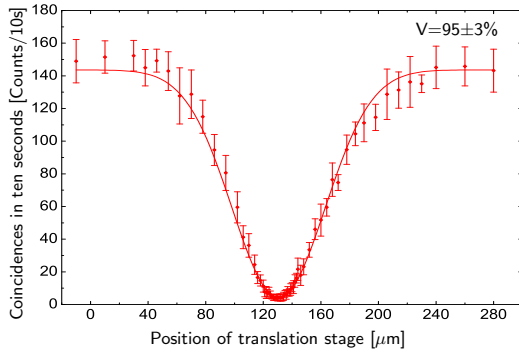


FIG. 5: An example of a HOM dip obtained in our experiment for the state $|\text{HH}\rangle\langle\text{HH}|$. The visibility, V , of the HOM dip above is $95 \pm 3\%$; other HOM dips observed for different polarization states are similar to this one.

temporal overlap between the photons, and the spatial overlap is controlled by adjusting the mirrors that direct the two beams to interfere on the BS. The interferometer is optimized at where the coincidence rate is minimal, as shown in Fig. 5. The visibility of the HOM dip is $95 \pm 3\%$. Imperfections in the BS ratios, the WPs, and the polarization controllers limit the maximum experimental achievable visibility of the dip. Nevertheless, the visibilities of the HOM dips obtained in our experiment exceed 90% for all the data collected for the construction of the histograms shown in Fig. 3.

-
- [1] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
 - [2] B. Terhal, Phys. Lett. A **271**, 319 (2000).
 - [3] O. Gühne and G. Tóth, Phys. Rep. **474**, 1 (2009).
 - [4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
 - [5] H. Zhu, Y. Teo, and B.-G. Englert, Phys. Rev. A **81**, 052339 (2010).
 - [6] M. Lewenstein, B. Kraus, J. Cirac, and P. Horodecki, Phys. Rev. A **62**, 052310 (2000).
 - [7] M. Kotowski, M. Kotowski, and M. Kuś, Phys. Rev. A **81**, 062318 (2010).
 - [8] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, G. M. D'Ariano, and C. Macchiavello, Phys. Rev. Lett. **91**, 227901 (2003).
 - [9] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
 - [10] Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháček, and Z. Hradil, Phys. Rev. Lett. **107**, 020404 (2011).
 - [11] Y. S. Teo, B. Stoklasa, B.-G. Englert, J. Řeháček, and Z. Hradil, Phys. Rev. A **85**, 042317 (2012).
 - [12] J. Řeháček and Z. Hradil, Phys. Rev. Lett. **90**, 127904 (2003).
 - [13] J. Dai, Y. L. Len, Y. S. Teo, L. A. Krivitsky, and B.-G. Englert, New J. Phys. **15**, 063011 (2013).
 - [14] P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum, and P. H. Eberhard, Phys. Rev. A **60**, R773 (1999).
 - [15] A. Ling, P. Han, A. Lamas-Linares, and C. Kurtsiefer, Laser Physics **16**, 1140 (2006).
 - [16] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).
 - [17] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Phys. Rev. A **61**, 010304 (2000).
 - [18] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky, Phys. Rev. A **75**, 042108 (2007).
 - [19] B.-G. Englert, C. Kurtsiefer, and H. Weinfurter, Phys. Rev. A **63**, 032303 (2001).